SMARTNav: A Visual-Inertial Dataset for Reliable Robotic State Estimation

Seyed Hojat Mirtajadini, Abeje Yenehun Mersha, Mohammad Aldibaja

Smart Mechatronics and Robotics (SMART) Research Group, Saxion University of Applied Science,
Enschede, The Netherlands – {s.h.mirtajadinigoki, a.y.mersha, m.a.j.aldibaja}@saxion.nl

Keywords: Visual-Inertial Odometry, SLAM, Dataset, Autonomous Navigation, LiDAR

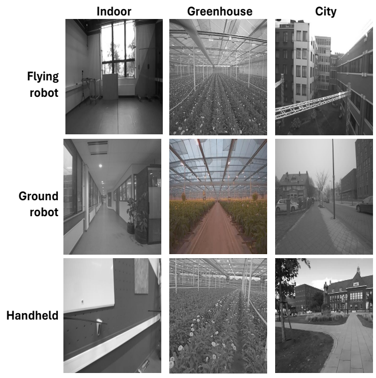
Abstract

Visual-inertial navigation has become a cornerstone for deploying robots in diverse environments. Despite significant progress, current approaches may easily fail to deliver reliable and robust navigation for industrial applications. Therefore, evaluating these methods using various datasets under challenging operational conditions is essential to ensure safe integration into robotic platforms. As such, this paper aims to enrich the availability of navigation datasets by introducing SMARTNav, which includes raw data obtained from stereo cameras and IMU sensors mounted on both ground and aerial robots. These robots were deployed in various operational scenarios across different environments, such as greenhouses, urban streets, indoor spaces, and near-building areas. The data includes challenges of navigating in GPS-denied areas, repetitive structures, featureless environments, and adverse lighting conditions. In order to provide corresponding ground-truth for each sequence, different techniques were deployed, such as Motion Capture System, Real Time Kinematics (RTK), and dense LiDAR-based Simultaneous Localization and Mapping (SLAM). Consequently, the resulting dataset can be used to address and validate key issues in vision-based state estimation, localization, and mapping for industrial applications. The SMARTNav dataset is accessible at: `https://saxionmechatronics.github.io/smartnav-dataset/`.

1. Introduction

Visual-Inertial Odometry (VIO) is a critical technology widely used in robotics and augmented reality to provide low-cost and high-frequency pose estimation, especially in environments with poor or unavailable GPS signals. This technology can be deployed in various indoor and outdoor environments for robot navigation. However, the performance of existing techniques is still not adequate for a reliable industrial application. Therefore, the availability of diverse datasets in terms of environment variety, sensor configurations, and addressing technical and operational challenges associated with Simultaneous Localization and Mapping (SLAM) is crucial before the integration of these methods into the robots' control loops. Current widely used datasets, such as EuRoC (Burri et al., 2016), TUM VI (Schubert et al., 2018), or KITTI (Geiger et al., 2013), cover mainly structured indoor environments or outdoor driving scenarios and lack the industrial robotic challenges faced when implementing VIO technology.

In this work, we present the SMARTNav dataset for evaluating VIO technologies in different industrial environments, covering various navigation scenarios using handheld, aerial, and ground robots. The environments include indoor corridors, outdoor near-building environments, streets, and greenhouses. As ground-truth position data, different measures were taken, including the use of motion capture systems, RTK, and dense LiDAR SLAM. The collected data are not only limited to camera and IMU (specific to VIO), but also high-resolution Lidar and GPS are recorded as well in long and flexible scenarios, mainly as supporting material to achieve the best estimate trajectory for accuracy assessment of VIO. Various challenges exist in different sequences of SMARTNav, caused by environmental characteristics, including repetitive textures and structures, featureless surfaces, and dynamic lighting conditions. Additionally, there are practical limitations such as vibrations, sample dropouts, and motion blur. Based on the above explanation, the main contributions of this paper are:



(a)

(b)

Figure 1. Overview of (a) diverse environments included in our dataset and (b) the ground and aerial robotic platforms used in our experiments.

https://doi.org/10.5194/isprs-annals-X-2-W2-2025-149-2025

Dataset	**Sensor Modality**				**Motion**			**Environment**				**Ground Truth**		
	Camera	IMU	LiDAR	Event Camera	Drone	Ground Vehicle	Handheld	Indoor	Streets	Near Buildings	Green-house	RTK	Dense SLAM/ Prior Maps	Motion Capture
KITTI (Geiger et al., 2013)	✓	✓	64-b	✗	✗	✓	✗	✗	✓	✓	✗	✓	✓	✗
EuRoC (Burri et al., 2016)	✓	✓	✗	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓
PennCOSYVIO (Pfrommer et al., 2017)	✓	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗	✗	✗	✗
TUM VI (Schubert et al., 2018)	✓	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗	✗	✗	✓
UZH-FPV (Delmerico et al., 2019)	✓	✓	✗	✓	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓
OpenLORIS-Scene (Shi et al., 2020)	✓	✓	2D	✗	✗	✓	✗	✓	✗	✗	✗	✗	✓	✓
Newer College Dataset (Ramezani et al., 2020)	✓	✓	64-b	✗	✗	✗	✓	✗	✗	✓	✗	✗	✓	✗
UMA-VI (Zuñiga-Noël et al., 2020)	✓	✓	✗	✗	✗	✗	✓	✓	✗	✓	✗	✗	✗	✓
M2DGR (Yin et al., 2021)	✓	✓	32-b	✓	✗	✓	✗	✓	✓	✓	✗	✓	✗	✓
NTU VIRAL (Nguyen et al., 2022)	✓	✓	64-b	✗	✓	✗	✗	✓	✗	✓	✗	✗	✗	✗
Hilti-Oxford (Zhang et al., 2022)	✓	✓	32-b	✗	✗	✗	✓	✓	✗	✓	✗	✗	✓	✗
M3ED (Chaney et al., 2023)	✓	✓	64-b	✓	✓	✓	✗	✗	✓	✓	✗	✓	✗	✓
SubT-MRS (Zhao et al., 2024)	✓	✓	16-b	✗	✓	✓	✓	✓	✗	✓	✗	✗	✓	✗
FusionPortableV2 (Wei et al., 2024)	✓	✓	128-b	✓	✗	✓	✓	✓	✓	✓	✗	✓	✗	✗
MARS-LVIG (Li et al., 2024)	✓	✓	limited FOV	✗	✓	✗	✗	✗	✗	✗	✗	✓	✗	✗
BotanicGarden (Liu et al., 2024)	✓	✓	16-b	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓	✗
MUN-FRL (Thalagala et al., 2024)	✓	✓	16-b	✗	✓	✗	✗	✗	✗	✓	✗	✓	✗	✗
SMARTNav (ours)	✓	✓	128-b	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

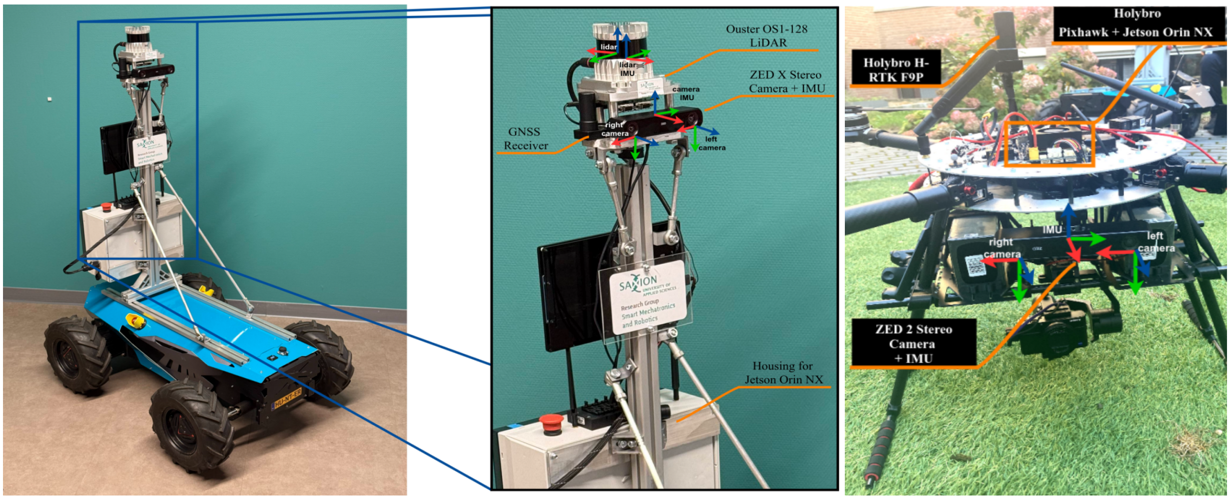
Table 1. Comparison of datasets in terms of used sensors, test platforms, environment variety, and ground-truth measurement. For LiDAR, letter b stands for the number of beams, determining the sensor's resolution.

- A diverse VIO dataset covering underrepresented and challenging environments for industrial applications of robots such as greenhouses, collected on different platforms.
- A multi-method approach for high-quality ground truth trajectory generation in diverse scenarios.

The rest of the manuscript is organized as follows. In section 2, we looked into the previously published SLAM and VIO datasets and highlight the novel aspects of our experiments. Also, section 3 details the collection, calibration, and ground truth measurement for our data sequences. Discussions and some insights into SMARTNav datasets are presented in section 4, and finally, section 5 brings a conclusion to the paper.

2. Related Works

Since the emergence of SLAM methods, various datasets have been published to introduce challenging use cases of these technologies in the real world. One of the pioneering and well-known datasets is KITTI (Geiger et al., 2013) that introduces a standard benchmark for developing GPS-denied navigation solutions primarily for automotive applications. Another publicly available dataset is PennCOSYVIO (Pfrommer et al., 2017), dedicated to the application of VIO in handheld devices, mainly in offices, labs, construction areas, and parking lots (outdoors). This dataset used fiducial markers as the source of ground truth position. As such, the dataset collectors could cover larger and



(a) (b)

Figure 2. Sensor configuration and experimental platforms a) VIDAR setup that is mounted on a wheeled robot for ground data sequences. b) SARAX (Alharbat et al., 2024) drone that is used for aerial sequences.

more arbitrary trajectories. TUM VI (Schubert et al., 2018) is another widely accepted benchmark for high-fidelity VIO sensor data, with hardware-level synchronized camera and IMU readings and 16-bit images beneficial for higher dynamic ranges. This data features both indoor and outdoor scenarios using a handheld setup; however, for the ground-truth, it relies on motion capture readings only at the beginning and end of the trajectories. UMA-VI (Zuñiga-Noël et al., 2020) exhibits a similar data collection strategy, however, it delivers a higher number of sequences with more challenging texture and illumination conditions. To focus more on the robotics aspect of visual-inertial localization, the EuRoC MAV (Burri et al., 2016) dataset was presented to provide synchronized image and IMU readings accompanied by a very accurate motion capture-based ground truth trajectory. Likewise, UZH-FPV (Delmerico et al., 2019) was published, collected from FPV drone flights; however, it emphasizes more on agile flights and drastic motions.

There is a large number of datasets that are not specifically collected for visual-inertial navigation and are extended to other sensor modalities as well. For example, OpenLORIS-Scene (Shi et al., 2020) is such a dataset that adds a 2D LiDAR sensor for ground robots' lifelong applications. Newer College Dataset (Ramezani et al., 2020) brings 3D LiDAR scans with 64 beams (vertical channels) to the recorded sequences, which extends the use-cases of the dataset to Lidar-Inertial Odometry (LIO) and LiDAR-based SLAM algorithms. A recent, yet similar in terms of platform and ground truth retrieval, is the Hilti-Oxford (Zhang et al., 2022). This dataset describes the use of dense prior maps of the environment and localizing LiDAR scans in the prior map, as a way to calculate the ground-truth trajectory. M2DGR (Yin et al., 2021) explores further indoor, street views, and near-building environments in the dataset, and depending on the availability, it uses motion capture or RTK as ground truth. A multi-sensor SLAM dataset for drone platforms is the NTU VIRAL (Nguyen et al., 2022), utilizing multiple cameras, IMUs, 2 LiDARs, and Ultra Wide Band (UWB), for both indoor and outdoor environments.

The recent datasets tend to broaden the test scenarios of the SLAM, both in terms of challenges, sensor modalities, used platforms, and experiment environments. M3ED (Chaney et al., 2023) is one of the first to collect data on multiple aerial and ground robots, and uses a motion capture system and precise GNSS for ground-truth depending on the test case. SubT-MRS (Zhao et al., 2024) is another multi-platform dataset with a wide range of sensors, captured in challenging multi-season environments, but more focused on subterranean environments.

One of the main barriers, limiting the complexity and length of the dataset sequences in the explored literature, was the choice of ground-truth measurement method. The datasets relying on motion capture systems are usually limited to small lab environments (Burri et al., 2016, Delmerico et al., 2019), or they cannot provide the ground truth for the entirety of the trajectory (Schubert et al., 2018, Zuñiga-Noël et al., 2020). GNSS-based methods are also limited to clear sky environments. However, recent datasets deploy dense prior maps of the intended environments and then localize the recorded LiDAR scans in the given accurate prior map (Zhao et al., 2024, Zhang et al., 2022, Ramezani et al., 2020). Alternatively, high resolution and accurate LiDAR SLAM was found as a sufficient solution for generating ground truth (Liu et al., 2024).

Building on the above overview, this paper seeks to enrich the pool of datasets available for validating visual-inertial odometry (VIO) methods in industrial settings. The SMARTNav dataset covers a wide range of challenging conditions, including GPS-denied areas, repetitive structures, featureless surfaces, and urban landscapes, using both aerial and ground robotic platforms. It also provides high-precision ground truth from high-resolution LiDAR, GNSS-RTK, and OptiTrack systems. Table 1 compares SMARTNav to the existing datasets discussed above, highlighting its unique contributions.

3. SMARTNav Dataset

3.1 Experimental Platforms

3.1.1 Ground Robot Fig. 2a shows the ground experimental platform with Visual-LiDAR (VIDAR) setup, GNSS receiver, and processing unit, which is a Jetson Orin NX. The LiDAR is an Ouster OS1-128 sensor, which provides 3D point clouds at 128×1024 resolution, 10 Hz frequency, 200 meters maximum range, and centimeter-level range accuracy. The LiDAR is also

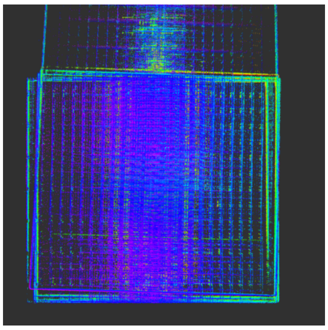
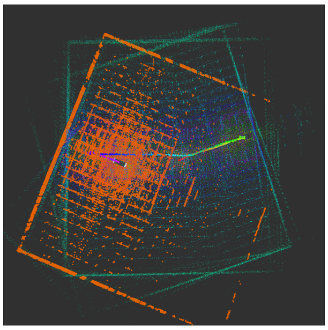
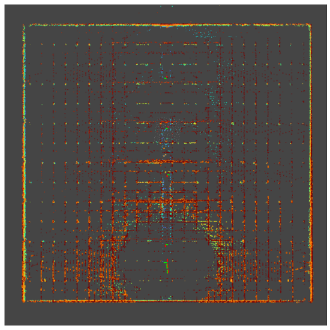
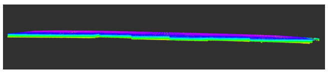
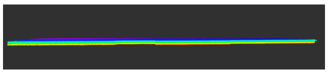
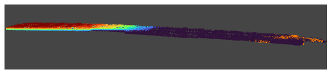
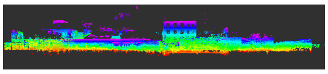
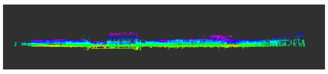
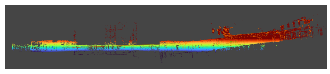
	FAST-LIO2 (Xu et al., 2022)	LeGO-LOAM (Shan and Englot, 2018)	GLIM (Koide et al., 2024)
`greenhouse_ground_1`			
`greenhouse_ground_2`			
`corridor_ground_1`			

Table 2. Comparison of LiDAR SLAM methods to provide the best estimate trajectory

equipped with an InvenSense ICM-20948 IMU sensor recorded at 100 Hz. Additionally, the ZED X stereo camera is used to record visual sensory data at different resolutions, color spaces, and frequencies depending on the testing environments. For the greenhouse tests, RGB images were recorded at 480×300 resolution and 15 Hz frequency. The reason for low resolution in these tests was using the wrong *Quality of Service* (QoS) policy of *best effort* for sensor messages in ROS 2, which prevented any buffering and caused message drops. Hence, we had to lower the sensor data resolution and rates to compensate for the effect. For the street-side walk, campus, and corridor tests, we fixed this issue by setting QoS to *reliable*. Then, grayscale images with 960×540 resolution were recorded with a consistent 15 Hz frequency.

3.1.2 Aerial Robot For the flying scenarios, we used different hardware setups due to modifications and upgrades. Initially, SARAX (Alharbat et al., 2024) drone platform (Fig. 2b) was used within a lab room equipped with an OptiTrack Prime 13 motion capture system. In these tests, the ZED 2 stereo camera was deployed, recording grayscale images with 672×376 resolution and 30 Hz rate. Furthermore, handheld experiments were conducted by carrying the drone around the room. We used the same platform and recording configuration in outdoor flight tests within our drone dome facility, an area surrounded by buildings. For greenhouse tests, we upgraded the SARAX platform camera with a ZED X camera, which has global shutter sensors and internal IMU dampers, reducing the vibration effects. Here, the images were recorded at 960×600 resolution and the rate was 30 Hz.

3.2 Ground-Truth

In the flight and handheld tests conducted in the lab room, the motion capture system provided the ground truth position data at above 180 Hz. In the outdoor flight test, we used a Holybro H-RTK F9P rover to measure RTK as the ground-truth data. The RTK data passed through the Extended Kalman Filter (EKF) from the PX4 autopilot, and the smoothed output of the autopilot has been recorded. The position variances from EKF were carefully checked for possible positioning inaccuracies, and only the trajectory segments with variances above 0.05 were accepted as ground-truth. In the drone dome tests, satellite signals for RTK were intermittently available due to proximity to buildings, yielding centimeter-level accuracy for parts of trajectories. In contrast, and to our surprise, the RTK provided strong satellite coverage with reasonable accuracy as ground truth for greenhouse flight tests.

In order to generate ground truth for the ground sequences, we investigated the performance of several SLAM-based methods. We checked the suitability of each method by visually inspecting the map consistency, optimized pose drifts in revisited areas, and correspondence with the paved trajectory on the satellite images.

Table 2 shows the comparison results between the three trajectory estimation methods, namely FAST-LIO2 (Xu et al., 2022), LeGO-LOAM (Shan and Englot, 2018), and GLIM (Koide et al., 2024). FAST-LIO2 did not have the mapping component of a typical SLAM algorithm, while the other two were capable of global pose-graph optimization to null the drifts in revisited places, i.e., loop closures. As clearly visible, for `greenhouse_ground_1`, GLIM has the best performance with no map drifts compared to the other methods. In `greenhouse_ground_2`, which is a long direct pavement, GLIM struggles to correctly follow the ground level, while LeGO-LOAM gives the most drift-free estimation, thanks to its ground-optimized odometry and mapping formulation. Moreover, Fig. 3a shows the trajectory outputs of FAST-LIO2, LeGO-LOAM, and a customer-grade GPS, drawn on a satellite map. By visual inspection of the visited locations on the map, and checking the accordance of start and end points, we concluded that LeGO-LOAM can provide accurate position estimation with less than 1 meter error.

3.3 Calibration

To measure each of the stereo cameras' intrinsic parameters, extrinsics between the two cameras, and extrinsics between the IMU and cameras, the Kalibr (Rehder et al., 2016) package was used. The manufacturer-reported values for these parameters are also present in the dataset; however, for IMU noise characteristics, we settled for the vendor's measured values.

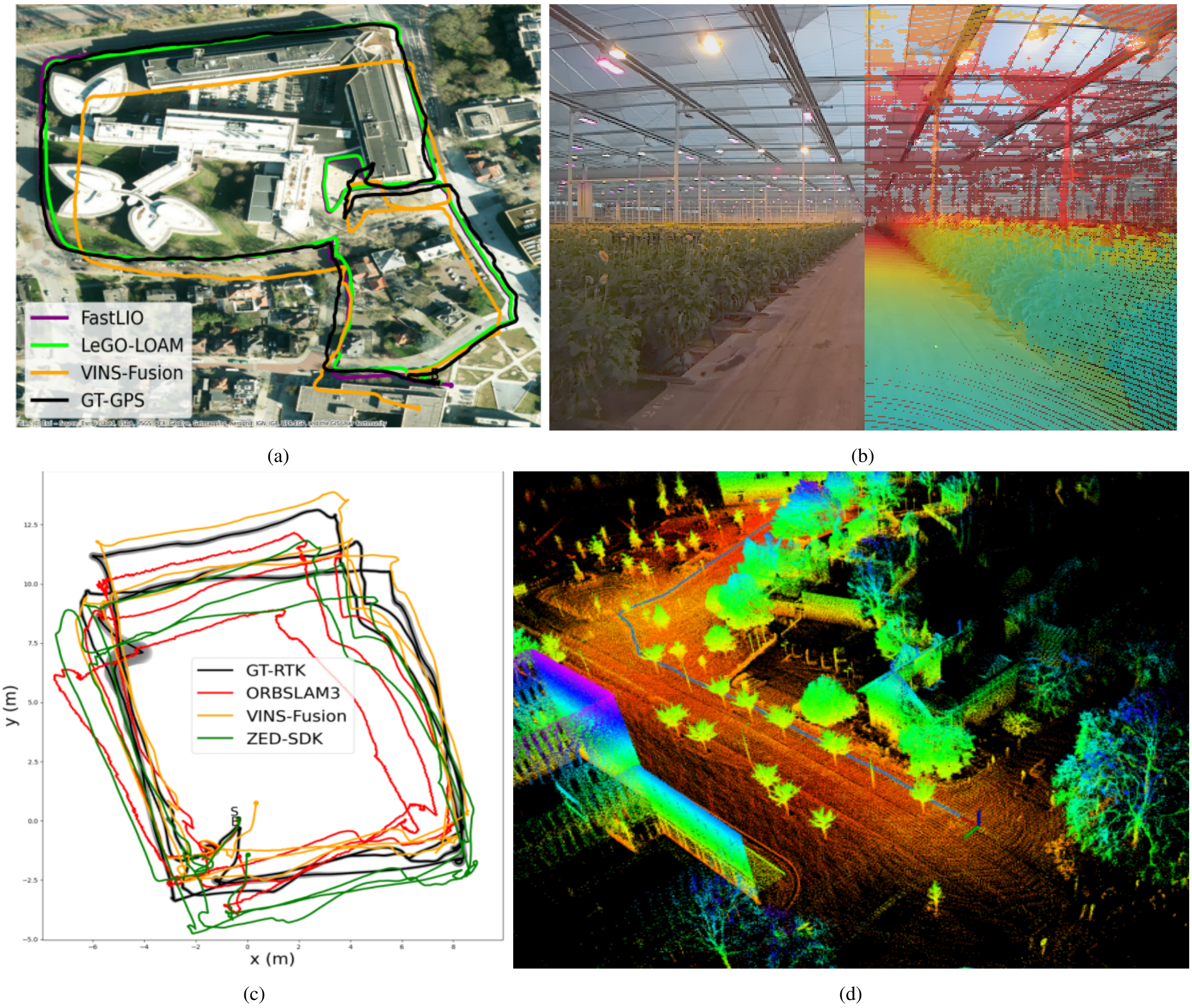


Figure 3. Some visualizations of the data in the SMARTNav dataset. a) Comparison between LiDAR-inertial odometry, visual-inertial odometry, and GNSS localization in a street test. b) Projection of LiDAR's dense point cloud on the image after camera-LiDAR calibration. c) Comparison plot of VIO methods in a drone flight test. d) The map that was created by accumulating the point clouds based on positions calculated by FAST-LIO2.

The `direct_visual_lidar_calibration` package (Koide et al., 2023) was used for targetless camera-LiDAR calibration. The method automatically and iteratively correlates the features in LiDAR point clouds to the visual representation in the camera image as demonstrated in Fig. 3b.

4. Discussion

Fig. 3 demonstrates some of the data modalities present in the SMARTNav dataset. In Fig. 3c, an initial assessment of some well-established VIO methods such as VINS-Fusion (Qin et al., 2018), ORBSLAM3 (Campos et al., 2021), and proprietary ZED-SDK, are compared to RTK ground-truth. For the RTK, the 5 times exaggerated covariance values are drawn in gray color around the trajectory. Since the data belongs to a flight test in a near-building environment, only at low covariance segments, the trajectory can serve as the ground truth. This behavior is observed in some other sequences in our dataset, mostly near trees and buildings with limited sky views.

The flight, ground, and handheld data collected in the greenhouse were challenging for VIO algorithms, due to the repetitive patterns of the plants. The `greenhouse_flight_4` data in Table 3 shows the feature tracking of VINS-Fusion that only relied on the ceiling patterns, consequently, the drift of VIO was considerable. Another frequent problem for the existing VIO algorithms, also reported in other benchmarks (Delmerico and Scaramuzza, 2018, Schubert et al., 2018), is the divergence of the state estimate, usually leading to erratic motions and unstable drifting, similar to `outdoor_flight_3` and `optitrack_flight_2` plots in Table 3. This behavior can be consequential for downstream control algorithms in robotic implementations. Providing a covariance estimation is a standard approach to avoid this issue, a requirement that most of the available optimization-based VIO methods do not have.

Similar to vision data, processing LiDAR data in the greenhouse test was challenging, not only due to the environmental structure but also because of data sample dropouts and relatively large motions between some consecutive raw point clouds.

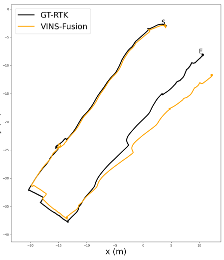
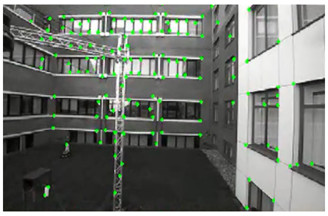
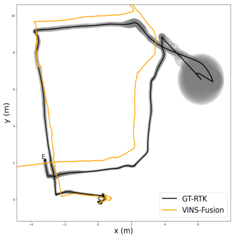
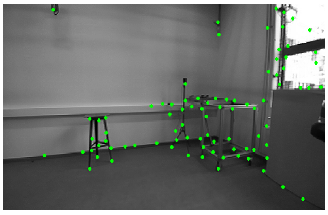
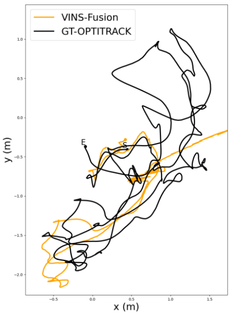
	Feature tracks	Trajectories	Test environment
greenhouse_flight_3			
outdoor_flight_3			
optitrack_flight_2			

Table 3. Analysis of VIO performance in our flight tests.

5. Conclusion

This paper described the SMARTNav dataset, gathered to assist in developing and evaluating vision-based navigation solutions. The data was collected over the span of multiple projects, using aerial, ground, and handheld motion platforms, in different indoor and outdoor spaces, most notably a flower greenhouse. We hope this data can help the further development of reliable visual-inertial navigation solutions for heterogeneous sensors and platforms.

Based on the analysis of the SMARTNav dataset, we identified several key challenges and further development areas. One of the pressing issues was the reliability of optimization and bundle-adjustment methods, and proper failure reporting through a calibrated covariance estimation. Additionally, partial discrepancies of individual sensors were observed in multiple instances in this data, requiring more attention to fusion methods. Particularly, using prior maps, created from dense LiDAR measurements and low-cost localization in these maps, is an appealing development area using our dataset. Despite the initial intent for studying VIO methods, many of the collected data can also be used for LiDAR-based SLAM development and evaluation. Lastly, expanding the current dataset by adding environments of robot implementation, a more diverse set of sensors, and mobile platforms is in our future work scope.

Acknowledgements

This work was partly supported by the Netherlands Organization for Scientific Research (NWO) via SIA RAAK-Public project (Van bestrijden naar beheersen van de EPR, No.10.015) and SIA RAAK-MKB project (Smart Greenhouses, No.17.014). The authors would like to thank the contributions of project managers Henri Huisman and Amin Zaami as well as Gerjen ter Maat for helping in the data collection.

References

Alharbat, A., Zwakenberg, D., Esmaeeli, H., Mersha, A., 2024. Sarax: An open-source software/hardware framework for aerial manipulators. *International Conference on Unmanned Aircraft Systems, ICUAS 2024*.

Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W., Siegwart, R., 2016. The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*, 35(10), 1157–1163.

Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M., Tardós, J. D., 2021. Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam. *IEEE transactions on robotics*, 37(6), 1874–1890.

Chaney, K., Cladera, F., Wang, Z., Bisulco, A., Hsieh, M. A., Korpela, C., Kumar, V., Taylor, C. J., Daniilidis, K., 2023. M3ed: Multi-robot, multi-sensor, multi-environment event dataset. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4016–4023.

Delmerico, J., Cieslewski, T., Rebecq, H., Faessler, M., Scaramuzza, D., 2019. Are we ready for autonomous drone racing? the uzh-fpv drone racing dataset. *2019 International Conference on Robotics and Automation (ICRA)*, IEEE, 6713–6719.

Delmerico, J., Scaramuzza, D., 2018. A benchmark comparison of monocular visual-inertial odometry algorithms for flying robots. *2018 IEEE international conference on robotics and automation (ICRA)*, IEEE, 2502–2509.

Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets robotics: The kitti dataset. *The international journal of robotics research*, 32(11), 1231–1237.

Koide, K., Oishi, S., Yokozuka, M., Banno, A., 2023. General, single-shot, target-less, and automatic lidar-camera extrinsic calibration toolbox. *2023 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 11301–11307.

Koide, K., Yokozuka, M., Oishi, S., Banno, A., 2024. Glim: 3d range-inertial localization and mapping with gpu-accelerated scan matching factors. *Robotics and Autonomous Systems*, 179, 104750.

Li, H., Zou, Y., Chen, N., Lin, J., Liu, X., Xu, W., Zheng, C., Li, R., He, D., Kong, F. et al., 2024. MARS-LVIG dataset: A multi-sensor aerial robots SLAM dataset for LiDAR-visual-inertial-GNSS fusion. *The International Journal of Robotics Research*, 43(8), 1114–1127.

Liu, Y., Fu, Y., Qin, M., Xu, Y., Xu, B., Chen, F., Goossens, B., Sun, P. Z., Yu, H., Liu, C. et al., 2024. Botanicgarden: A high-quality dataset for robot navigation in unstructured natural environments. *IEEE Robotics and Automation Letters*, 9(3), 2798–2805.

Nguyen, T.-M., Yuan, S., Cao, M., Lyu, Y., Nguyen, T. H., Xie, L., 2022. Ntu viral: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint. *The International Journal of Robotics Research*, 41(3), 270–280.

Pfrommer, B., Sanket, N., Daniilidis, K., Cleveland, J., 2017. Penncosyvio: A challenging visual inertial odometry benchmark. *2017 IEEE International Conference on Robotics and Automation, ICRA 2017, Singapore, Singapore, May 29 - June 3, 2017*, 3847–3854.

Qin, T., Li, P., Shen, S., 2018. Vins-mono: A robust and versatile monocular visual-inertial state estimator. *IEEE transactions on robotics*, 34(4), 1004–1020.

Ramezani, M., Wang, Y., Camurri, M., Wisth, D., Mattamala, M., Fallon, M., 2020. The newer college dataset: Handheld lidar, inertial and vision with ground truth. *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 4353–4360.

Rehder, J., Nikolic, J., Schneider, T., Hinzmann, T., Siegwart, R., 2016. Extending kalibr: Calibrating the extrinsics of multiple imus and of individual axes. *2016 IEEE International Conference on Robotics and Automation (ICRA)*, IEEE, 4304–4311.

Schubert, D., Goll, T., Demmel, N., Usenko, V., Stückler, J., Cremers, D., 2018. The tum vi benchmark for evaluating visual-inertial odometry. *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 1680–1687.

Shan, T., Englot, B., 2018. Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, IEEE, 4758–4765.

Shi, X., Li, D., Zhao, P., Tian, Q., Tian, Y., Long, Q., Zhu, C., Song, J., Qiao, F., Song, L. et al., 2020. Are we ready for service robots? the openloris-scene datasets for lifelong slam. *2020 IEEE international conference on robotics and automation (ICRA)*, IEEE, 3139–3145.

Thalagala, R. G., De Silva, O., Jayasiri, A., Gubbels, A., Mann, G. K., Gosine, R. G., 2024. MUN-FRL: A Visual-Inertial-LiDAR Dataset for Aerial Autonomous Navigation and Mapping. *The International Journal of Robotics Research*, 43(12), 1853–1866.

Wei, H., Jiao, J., Hu, X., Yu, J., Xie, X., Wu, J., Zhu, Y., Liu, Y., Wang, L., Liu, M., 2024. Fusionportablev2: A unified multi-sensor dataset for generalized slam across diverse platforms and scalable environments. *The International Journal of Robotics Research*, 02783649241303525.

Xu, W., Cai, Y., He, D., Lin, J., Zhang, F., 2022. Fast-lio2: Fast direct lidar-inertial odometry. *IEEE Transactions on Robotics*, 38(4), 2053–2073.

Yin, J., Li, A., Li, T., Yu, W., Zou, D., 2021. M2dgr: A multi-sensor and multi-scenario slam dataset for ground robots. *IEEE Robotics and Automation Letters*, 7(2), 2266–2273.

Zhang, L., Helmberger, M., Fu, L. F. T., Wisth, D., Camurri, M., Scaramuzza, D., Fallon, M., 2022. Hilti-oxford dataset: A millimeter-accurate benchmark for simultaneous localization and mapping. *IEEE Robotics and Automation Letters*, 8(1), 408–415.

Zhao, S., Gao, Y., Wu, T., Singh, D., Jiang, R., Sun, H., Sarawata, M., Qiu, Y., Whittaker, W., Higgins, I. et al., 2024. Subt-mrs dataset: Pushing slam towards all-weather environments. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 22647–22657.

Zuñiga-Noël, D., Jaenal, A., Gomez-Ojeda, R., Gonzalez-Jimenez, J., 2020. The UMA-VI dataset: Visual–inertial odometry in low-textured and dynamic illumination environments. *The International Journal of Robotics Research*, 39(9), 1052-1060. https://doi.org/10.1177/0278364920938439.